

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

~~CLAIMS~~

[Claim(s)]

[Claim 1] the environmental sound analysis apparatus which extracts the time section when a specific sound is contained [from] among environmental sound signals -- it being and with a frame division means to generate two or more time frames which divided said environmental sound signal into the signal for every predetermined time interval A parameter calculation means to compute the description parameter value of the acoustical property which was generated with this frame division means and which was beforehand defined for said every time frame, The environmental sound analysis apparatus characterized by having a frequency-distribution function decision means to determine the frequency-distribution function showing the frequency distribution of said description parameter value computed with this parameter calculation means, and a peak detection means to detect the peak part of the frequency-distribution function determined with this frequency-distribution function decision means.

[Claim 2] The A/D converter with which said frame division means carries out A/D conversion of said environmental sound signal in an environmental sound analysis apparatus according to claim 1, The frame memory which classified said two or more time frames into two or more fields memorized, respectively, The environmental sound analysis apparatus characterized by having the division control means which makes each field of said frame memory memorize the conversion result of said A/D converter as said time frame for every part corresponding to the specific time die length which continued in time.

[Claim 3] The environmental sound analysis apparatus characterized by having had a means by which said parameter calculation means computed the average strength of a signal for said every time frame, in the environmental sound analysis apparatus according to claim 1 or 2, and having a means to determine the frequency-distribution function with which said frequency-distribution function decision means is expressed with the number of said time frames corresponding to the numeric value of the average strength of said signal.

[Claim 4] The environmental sound analysis apparatus characterized by having the weighting means which carries out weighting based on a frequency to the signal of said time frame generated with said frame division means in the preceding paragraph of said parameter calculation means in an environmental sound analysis apparatus according to claim 1 to 3.

[Claim 5] The environmental sound analysis apparatus characterized by having had a means by which said parameter calculation means computed the peak factor of a signal for said every time frame, in the environmental sound analysis apparatus according to claim 1 or 2, and having a means to determine the frequency-distribution function with which said frequency-distribution function decision means is expressed with the number of said time frames corresponding to the numeric value of the peak factor of said signal.

[Claim 6] The environmental sound analysis apparatus characterized by having a frame extract means to extract only the time frame corresponding to the peak part detected with said peak detection means in an environmental sound analysis apparatus according to claim 1 to 5.

[Claim 7] The environmental sound analysis apparatus characterized by having a means to judge

whether the location of the peak part of the frequency-distribution function which said frequency-distribution function decision means determined is clear, and a means to change the total of said time frame used according to the judgment result of this means, in an environmental sound analysis apparatus according to claim 1 to 6.

.....
[Translation done.]

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-298698
(43)Date of publication of application : 12.11.1996

(51)Int.Cl. H04R 25/00

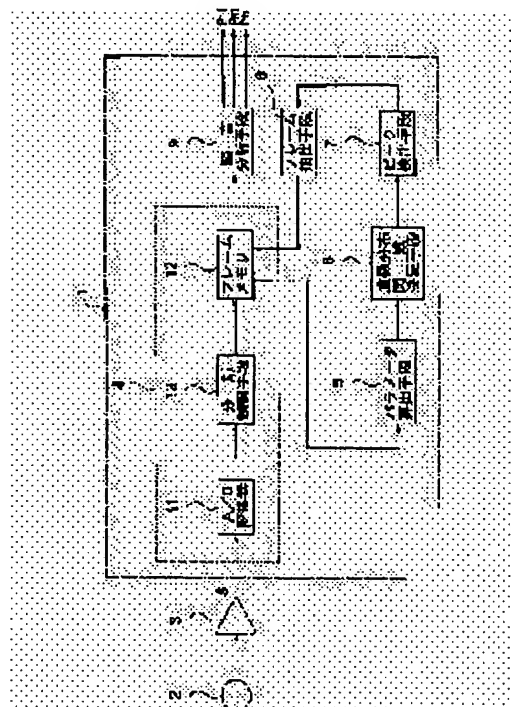
(21)Application number : 07-103449 (71)Applicant : RION CO LTD
(22)Date of filing : 27.04.1995 (72)Inventor : SAKAMOTO SHINICHI
OISHI TOMOKO

(54) ENVIRONMENTAL SOUND ANALYZER

(57)Abstract:

PURPOSE: To highly accurately extract a noise section including no conversational sound out of an environmental sound by determining a function expressing the frequency distribution of a prescribed acoustic feature parameter value in each plural time frames generated by dividing the environmental sound and detecting the peak part of the frequency distribution function.

CONSTITUTION: In this environmental sound analyzer, a dividing means 4 divides an environmental sound S so that prescribed time corresponds to one frame and a frequency distribution function determining means 6 determines a frequency distribution function for all frequency based upon the average power level P of each frame found by a parameter calculating means 5. A peak detecting means 7 regards a frame having the average power level P_x of a peak detected based upon the distribution function as the frame corresponding to an environmental noise, a frame extracting means 8 reads out the frame having the P_x from a frame memory 12 and sends the read frame to an analyzing means 9. Consequently only a signal in a noise section including no voice can be highly accurately extracted and an output sound to be easily listened can be obtained.



LEGAL STATUS

[Date of request for examination]	28.11.1997
[Date of sending the examiner's decision of rejection]	
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]	
[Date of final disposal for application]	
[Patent number]	2905112
[Date of registration]	26.03.1999
[Number of appeal against examiner's decision of rejection]	
[Date of requesting appeal against examiner's decision of rejection]	
[Date of extinction of right]	

Copyright (C); 1998,2003 Japan Patent Office

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates to the environmental sound analysis apparatus for extracting the time sections containing a specific sound, such as the section (silent section which does not contain a conversation sound) of the noise included in an environmental sound, and/or the section (voice section) of a conversation sound, concerning an environmental sound analysis apparatus.

[0002]

[Description of the Prior Art] Since the environmental level or the environmental (operating environment) property of the noise which uses a mobile equipment is not fixed when using the device (a "mobile equipment" is called below.) which service spaces, such as a headphone stereo cassette tape recorder, hearing aid, a car stereo, a portable type telephone, and a walkie-talkie, do not specify generally, changing the acoustical property of a mobile equipment automatically according to an operating environment is performed.

[0003] If it does not change tone quality in coming to sense noisy and using it further in locations where noise level is still higher, such as inside of a subway, if it does not make sound volume low in using it to raising sound volume in the location where noise level, such as bustle, is high since the level of a conversation sound also becomes low in hearing aid in using it in quiet locations of Nighttime, such as domestic, it becomes impossible for example, to catch.

[0004] Then, in order to adjust the acoustical property of hearing aid automatically according to change of such an operating environment for example, the thing (the so-called ANS --) to which it was made to change automatically frequency characteristics (tone quality) and sound volume according to the sound pressure level of an input sound (environmental sound) or the sound pressure level, and its persistence time to ** hearing aid Refer to JP,52-50646,B and U.S. Pat. No. 4,025,721, ** The thing (refer to the so-called K amplifier and U.S. Pat. No. 5,131,046) to which it was made to change the gain of a high frequency band automatically according to the sound pressure level of the high frequency band of the input sound (environmental sound) to hearing aid like ANS is known.

[0005] Moreover, in hearing aid, in order that a user may make a conversation sound easy to catch, it is made to perform various signal processing to an input sound. For example, by performing analysis composition processing to an input sound, and changing a frequency to the lower one paying attention to being hard to catch a sound with a high frequency in many hearing impairment people A hearing impairment person tends to catch a conversation sound, and make him become, or What makes it easy to catch also to elderly people is performed by detecting, the break, i.e., the silent section, of a sentence of a conversation sound, securing time allowances by packing the time amount, and carrying out speech rate transform processing which makes late the conversation speed of sound outputted from hearing aid.

[0006] By the way, although it does not produce un-arranging so much even if it changes sound volume, tone quality, etc. of an output sound according to all environmental sound signals, if it is in a headphone stereo cassette tape recorder, a car stereo, etc. aiming at conversation in a mobile equipment which was mentioned above If it is in the hearing aid aiming at conversation, a portable type telephone, a walkie-

talkie, etc., it is on the contrary hard coming to catch a conversation sound in having changed the acoustical property of an output sound according to all environmental sound signals. If a loudness level of sound is lowered according to all the inputted environmental sound signals, it becomes low to the level of not only the noise but a conversation sound, and it becomes impossible that is, to catch a conversation sound, since the conversation sound which contains its voice in an environmental sound is contained. [not only the noise but]

[0007] Moreover, it is on the contrary hard coming to catch a conversation sound in having performed necessary signal processing to all environmental sound signals with the hearing aid which performs signal processing in order to make a conversation sound easy to catch. For example, if the frequency of the whole input sound is changed to the lower one, it changes also except a conversation sound, and a remarkable unnatural sound will hear or it will be hard coming to carry out identification of the sound source of an environmental sound. Moreover, although expanding processing of the environmental sound signal inputted must be carried out and it must once store in memory etc. in order to carry out speech rate transform processing, when it elongates to the signal of the noise, an output sound becomes redundancy, it becomes impossible to catch, and the memory space to need will increase.

[0008] Then, it discriminates from the noise and a conversation sound from among environmental sounds, and only the noise section or the voice section (section containing a conversation sound) is extracted from an environmental sound signal, and it is necessary to adjust an acoustical property or to perform necessary signal processing. Therefore, the section, the voice section, etc. of the noise which are conventionally included in an environmental sound, various things are proposed as an environmental sound analysis apparatus for extracting the time section containing a specific sound (JP,6-93199,B --) By comparing with the reference level as which, as for these, reference, such as JP,6-32001,B, determined the level of an environmental sound beforehand fundamentally, beyond over fixed time amount, for example, environmental sound level distinguishes the part at the time of below fixed level as it is the noise section.

[0009]

[Problem(s) to be Solved by the Invention] However, if it is in an environmental sound analysis apparatus [reference level / level / of an environmental sound] in order to extract the noise section and/or the voice section from an environmental sound signal, as mentioned above In noise level is high and being close to conversation sound level The section (noise section) when a conversation sound does not exist since an environmental sound does not become below fixed level even when a conversation sound does not exist is undetectable. Conversely, when conversation sound level is low the conversation sound exists -- environmental sound level -- below fixed -- becoming -- a conversation sound -- not existing (it being the noise section) -- the technical problem that the analysis precision of detect of an environmental sound is bad occurs.

[0010] This invention is made in view of the above-mentioned point, and it aims at offering the environmental sound analysis apparatus which can extract the time section when a specific sound is contained [from] among environmental sound signals with high precision.

[0011]

[Means for Solving the Problem] In order to solve the above-mentioned technical problem the environmental sound analysis apparatus of claim 1 the environmental sound analysis apparatus which extracts the time section when a specific sound is contained [from] among environmental sound signals -- it being and with a frame division means to generate two or more time frames which divided said environmental sound signal into the signal for every predetermined time interval A parameter calculation means to compute the description parameter value of the acoustical property which was generated with this frame division means and which was beforehand defined for said every time frame, It had a frequency-distribution function decision means to determine the frequency-distribution function showing the frequency distribution of said description parameter value computed with this parameter calculation means, and a peak detection means to detect the peak part of the frequency-distribution function determined with this frequency-distribution function decision means.

[0012] The environmental sound analysis apparatus of claim 2 is set to the environmental sound analysis

apparatus of above-mentioned claim 1. The A/D converter with which said frame division means carries out A/D conversion of said environmental sound signal, The frame memory which classified said two or more time frames into two or more fields memorized, respectively, It considered as the configuration equipped with the division control means which makes each field of said frame memory memorize the conversion result of said A/D converter as said time frame for every part corresponding to the specific time die length which continued in time.

[0013] The environmental sound analysis apparatus of claim 3 was taken as the configuration which was equipped with a means by which said parameter calculation means computes the average strength of a signal for said every time frame, and is equipped with a means to determine the frequency-distribution function with which said frequency-distribution function decision means is expressed with the number of said time frames corresponding to the numeric value of the average strength of said signal in above-mentioned claim 1 or the environmental sound analysis apparatus of 2.

[0014] The environmental sound analysis apparatus of claim 4 was equipped with the weighting means which carries out weighting based on a frequency to the signal of said time frame generated with said frame division means in the preceding paragraph of said parameter calculation means in above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 3.

[0015] In claim 1 or the environmental sound analysis apparatus of 2, the environmental sound analysis apparatus of claim 5 was equipped with a means by which said parameter calculation means computes the peak factor of a signal for said every time frame, and was taken as the configuration which said frequency-distribution function decision means equips with a means to determine the frequency-distribution function expressed with the number of said time frames corresponding to the numeric value of the peak factor of said signal.

[0016] The environmental sound analysis apparatus of claim 6 was equipped with a frame extract means to extract only the time frame corresponding to the peak part detected with said peak detection means, in above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 5.

[0017] The environmental sound analysis apparatus of claim 7 was equipped with a means to judge whether the location of the peak part of the frequency-distribution function which said frequency-distribution function decision means determined is clear, and a means to change the total of said time frame used according to the judgment result of this means, in above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 6.

[0018]

[Function] The environmental sound analysis apparatus of claim 1 generates two or more time frames which divided the environmental sound signal into the signal for every predetermined time interval with the frame division means. The description parameter value of the acoustical property which was generated with the parameter calculation means and which was beforehand defined for every time frame is computed. By determining the frequency-distribution function which expresses the frequency distribution of the description parameter value with a frequency-distribution function decision means, and detecting the peak part of a frequency-distribution function with a peak detection means The voice section containing the time section which contains [from] a specific sound among environmental sounds, for example, the noise section which does not contain a conversation sound from the inside of an environmental sound, and a conversation sound can be extracted with high precision.

[0019] The environmental sound analysis apparatus of claim 2 is set to the environmental sound analysis apparatus of above-mentioned claim 1. The A/D converter with which a frame division means carries out A/D conversion of the environmental sound signal, The frame memory which classified two or more time frames into two or more fields memorized, respectively, Since it has the division control means which makes each field of a frame memory memorize the conversion result of an A/D converter as a time frame for every part corresponding to the specific time die length which continued in time, division into the time frame of an environmental sound signal can be performed easily at high speed.

[0020] Since the environmental sound analysis apparatus of claim 3 was equipped with a means to by which a parameter calculation means computes the average strength of a signal for every time frame and is equipped with a means determine the frequency-distribution function with which a frequency-

distribution function decision means is expressed with the number of the time frames corresponding to the numeric value of the average strength of a signal, it can extract the time section which contains [from] a specific sound among environmental sounds with high precision in above-mentioned claim 1 or the environmental sound analysis apparatus of 2.

[0021] In above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 3, since the environmental sound analysis apparatus of claim 4 was equipped with the weighting means which carries out weighting based on a frequency to the signal of the time frame generated with the frame division means in the preceding paragraph of a parameter calculation means, it can extract the time section which contains [from] a specific sound among environmental sounds much more with high precision.

[0022] In claim 1 or the environmental sound analysis apparatus of 2, since it had a means by which a parameter calculation means computed the peak factor of a signal for every time frame and the frequency-distribution function decision means is equipped with a means to determine the frequency-distribution function expressed with the number of the time frames corresponding to the numeric value of the peak factor of a signal, the environmental sound analysis apparatus of claim 5 can extract the time section which contains [from] a specific sound among environmental sounds with high precision.

[0023] In above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 5, since the environmental sound analysis apparatus of claim 6 was equipped with a frame extract means to extract only the time frame corresponding to the peak part detected with the peak detection means, it can extract only the time section (signal part) which contains [from] a specific sound among environmental sound signals.

[0024] The environmental sound analysis apparatus of claim 7 is set to above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 6. Since the frequency-distribution function decision means was equipped with a means to judge whether the location of the peak part of the determined frequency-distribution function is clear, and a means to change the total of the time frame used according to the judgment result of this means By the case where noise level is changed etc., when the location of the peak part of a frequency-distribution function is indefinite, a peak part can be detected more to high degree of accuracy by making [many] the total of the time frame to be used.

[0025]

[Example] Hereafter, the example of this invention is explained based on an accompanying drawing. Drawing 1 is the block diagram showing one example of the environmental sound analysis apparatus concerning this invention.

[0026] This environmental sound analysis apparatus 1 inputs the signal which amplified the microphone signal from the microphone 2 which collects an environmental sound with amplifier 3 as an environmental sound signal S. A frame division means 4 to generate two or more time frames which divided the environmental sound signal S into the signal for every predetermined time interval, A parameter calculation means 5 to compute the description parameter value of the acoustical property which was generated with this frame division means 4 and which was beforehand defined for every time frame, A frequency-distribution function decision means 6 to determine the frequency-distribution function showing the frequency distribution of the description parameter value computed with this parameter calculation means 5, A peak detection means 7 to detect the peak part of the frequency-distribution function determined with this frequency-distribution function means 6, It has a frame extract means 8 to extract only the time frame corresponding to the peak part detected with this peak detection means 7, and an ambient noise analysis means 9 to analyze an ambient noise based on the time frame extracted with this frame extract means 8.

[0027] The frame division means 4 carries out A/D conversion of the environmental sound signal S, and consists of A/D converter 11 which carries out digital coding, a frame memory 12 which classified two or more time frames into two or more fields memorized, respectively, and a division control means 13 which makes each field of a frame memory 12 memorize the conversion result of A/D converter 11 as a time frame for every part corresponding to the specific time die length which continued in time.

[0028] The parameter calculation means 5 reads the time frame (environmental sound signal between

specific time die length) memorized to each field of the frame memory 12 of the frame division means 4 one by one, and computes (average power level) as description parameter value in average strength for every time frame. The frequency-distribution function decision means 6 determines the frequency-distribution function showing the number of the time frames corresponding to the numeric value of the average power level of the time frame which the parameter calculation means 5 computed (frequency).

[0029] The peak detection means 7 detects the peak value which appears in the field where the number (frequency-distribution function) Kamihira ** power level of time frames corresponding to the numeric value of the average power level which the frequency-distribution function decision means 6 determined is low. The frame extract means 8 extracts from from only the time frame which has the average power level corresponding to the peak value detected with the peak detection means 7 among each time frame memorized by the frame memory 12, and is made to read it to the noise analysis means 9.

[0030] After the noise analysis means 9 carries out frequency analysis of the time frame (environmental sound signal of the specific time section) read from the frame memory 12 by the FFT algorithm, it computes each band power of a low frequency band, an inside frequency band, and a high-frequency band, and outputs the low frequency band power signal Pl, the inside frequency band power signal Pm, and the high-frequency band power signal Ph.

[0031] An operation of the example constituted as mentioned above is explained also with reference to drawing 2 thru/or drawing 5. If here explains the outline of the environmental sound analysis by this invention, it notes that the acoustical property of the sound which comes from the specific sound source in the scene of specification [this invention] is fixed in general. That is, since the description parameter value of acoustical properties, such as the level and frequency spectrum of the noise, and a peak factor, changes with the distance from the strength of a noise source, or a noise source, it cannot identify it with the absolute level. Moreover, the absolute level changes with an utterance sound, the reinforcement of utterance, and the distance from an utterance person also about a conversation sound.

[0032] However, in a specific scene, since the sound from a specific sound source is maintaining fixed level and a fixed acoustical property if a situation does not change, when it divides an environmental sound signal for every short time amount and asks for the frequency-distribution function of the description parameter value of the acoustical property in each time amount, in a frequency-distribution function, two or more peaks which embraced the class of sound will make it. For example, although a conversation sound is emitted on level somewhat higher than the noise when talking under the steady noise, as shown in a conversation sound at drawing 3, the silent section NB which is the section when a conversation sound is not surely contained exists, and although it is very a short time, since this silent section NB is the moment voice influence disappears, when an ambient noise exists in a perimeter, it turns into the section which the ambient noise itself appears.

[0033] At this time, what hits during conversation sound utterance among the environmental sound signals divided for every short time amount will have the average power corresponding to the level of a conversation sound, and what hits when it is not under utterance will have the average power corresponding to noise level. Therefore, two peaks will appear in a frequency-distribution function, the peak with larger average power will be equivalent to the signal at the time of being under utterance, and the peak of the smaller one will be equivalent to the signal at the time only of the noise which is not under utterance.

[0034] Then, what is necessary will be to perform signal-processing processing only to a signal with the average power of the peak of the larger one, performing the above-mentioned processing continuously that what is necessary is to collect only things with the average power equivalent to the peak of the smaller one of the divided environmental sound signals, and just to analyze if only the noise is analyzed, if signal-processing processing is performed only to a conversation sound.

[0035] Thus, even when the level of a conversation sound or the noise differs, by detecting the peak part of a frequency-distribution function, the time section containing a desired sound can be extracted with high precision, and necessary signal processing and adjustment of an acoustical property can be performed only in the required range.

[0036] When this example which analyzes an environmental sound below, detects the noise section and

analyzes the acoustical property is explained concretely, the environmental sound analysis apparatus 1 the counter a for carrying out counting of the number of time frames which created frequency distribution as shown in drawing 2 -- resetting ($a = 0$) -- the total number of time frames (this is called "the total frequency".) which is the need at creation of frequency distribution and which was defined beforehand It sets to Counter z (the $z =$ total frequency). And A/D conversion of the environmental sound signal S which amplified the microphone signal from a microphone 1 with amplifier 2 is carried out with A/D converter 11 of the frame division means 4 for every fixed time amount, and it changes into the data stream of a digital sign.

[0037] Then, a break S, i.e., an environmental sound signal, is divided into the signal for every predetermined time interval with the number equivalent to the die length of the short time amount LD (sec) which defined the data stream of an environmental sound signal beforehand by the division control means 13 of the frame division means 4. The part (signal = data stream) corresponding to the die length of the time amount LD (sec) which continued in time is generated as one time frame, and the sequential storage of each generated time frame is carried out for every frame to each field which the frame memory 12 divided beforehand. namely, the signal part of the die length of the time amount LD (sec) short about the environmental sound signal S (for actually processing, it is a data stream after A/D conversion) inputted as shown in drawing 4 -- as one time frame F -- time frames F1, F2, and F3 as -- it stores in each field of a frame memory 12. In addition, although the initiation timing of each time frame F is overlapped in the example of this drawing, this is for improving precision.

[0038] And the average power level P is computed for every time frame with the parameter calculation means 5 by reading the data stream of the environmental sound signal corresponding to the die length of the predetermined time amount LD memorized to each field of a frame memory 12 for every frame during one time. This average power level P can be computed by calculating $P = (x_0^2 + x_1^2 + x_2^2 + \dots + x_n^2) / n$, when the value acquired by f (Hz) and "LDxf" in a sampling frequency is set to n (measurement size).

[0039] Subsequently, based on the average power level P of the time frame computed with the parameter calculation means 5, a creation decision of the frequency-distribution function with which it is expressed as the number of the time frames corresponding to the numeric value of the average power level P is made with the frequency function decision means 6. That is, as shown in drawing 5, the frequency-distribution function which makes an axis of abscissa the average power level P, and makes an axis of ordinate the number of the time frames which have the average power level P concerned is created. This can form the counter corresponding to the average power level P of the time frame which the parameter calculation means 5 computed, and can obtain it by carrying out counting of the number of the time frames which have the average power level P concerned by carrying out the increment (+1) of the counter which corresponds whenever the power level P concerned is computed.

[0040] And by distinguishing whether after creating the frequency distribution about the time frame concerned read from the frame memory 12, the increment (+1) of the counter a was carried out, the value of Counters a and z was compared, and it became $a > z$, it judges whether the frequency distribution for the total frequency (the total number of time frames) defined beforehand have been created, and the above processings are repeated until it can create the frequency distribution for the total frequency.

[0041] Thus, a frequency-distribution function as shown in drawing 5 is obtained by creating the frequency distribution (function of the number of average power level pair time frames) of each average power level P about the time frame for the total frequency. namely, when the sound signal by the conversation sound mixes in an environmental sound signal Although the average power level P of the time frame which voice mixes becomes high, and it will become distribution as shown by the dotted line if crest type distribution as shown in the range of A in this drawing is shown and there is no ambient noise in the range of B of this drawing at this time When an ambient noise exists, as a continuous line shows, it is another crest (this is called "the 1st peak"). It appears.

[0042] Then, the peak detection means 7 detects the 1st peak (peak part of a frequency-distribution function) of the above from the frequency-distribution function created with the frequency-distribution function decision means 6, and it is presumed that the time frame which has the average power level Px

of this 1st peak is a time frame corresponding to an ambient noise. By it, the frame extract means 8 makes the time frame which has the average power level P_x , i.e., the time frame corresponding to an ambient noise, read from among each time frame memorized by the frame memory 12, and is sent out for the noise analysis means 9. At this time, the frequency-distribution function decision means 6 clears the created frequency-distribution function.

[0043] In this case, when there are few time frames which have the average power level P_x and it is not suitable for analysis, it judges whether the location of the 1st peak is indefinite, and when the location of the 1st peak is indefinite, analysis precision can be secured by sending out the time frame which has the average power level before and behind the 1st peak for the noise analysis means 9. moreover, when such, it adjusts in the direction which increases the total frequency z (total of the time frame used for the decision of a frequency-distribution function) automatically -- if it is made like (z is changed), the precision of a frequency-distribution function can improve and analysis precision can also be raised further.

[0044] With the noise analysis means 9 which received the time frame corresponding to an ambient noise as mentioned above, after carrying out frequency analysis of the data stream of the time frame inputted by the FFT algorithm, each band power of a low frequency band, an inside frequency band, and a high-frequency band is computed, and the low frequency band power signal P_l , the inside frequency band power signal P_m , and the high-frequency band power signal P_h are outputted.

[0045] Thus, two or more time frames which come to divide an environmental sound signal into the signal for every short time interval are generated. Compute the average power level (average strength) for every time frame as description parameter value, and by determining a frequency-distribution function based on the computed description parameter value, and detecting the peak part of this frequency-distribution function Only the signal of the time section (noise section) containing a necessary sound (the environmental sound which does not contain voice in this example = noise) can be extracted with high precision. The hearing aid sound of hearing aid can be adjusted now the optimal by adjusting acoustical properties, such as sound volume of hearing aid, and tone quality, using the low frequency band power signal P_l which followed, for example, was acquired as mentioned above, the inside frequency band power signal P_m , and the high-frequency band power signal P_h .

[0046] Next, the speech processing which applied this invention is explained with reference to drawing 6. This speech processing is made to perform signal processing only to parts other than the noise section of the environmental sound signals. First, while resetting the counter x for storing the average power level P_x of the time frame corresponding to the noise section ($x=0$) After setting the total frequency which reset the counter a for the numbers of time frames like the above, and was beforehand set to the total counter z for frequencies, Sequential storing of the time frames DATA0-n which divided and generated similarly the environmental sound signal which amplified the microphone signal from a microphone with amplifier is carried out with the above-mentioned example having explained at a frame memory. The data stream of the time frames DATA0-n memorized to each field is read from a frame memory for every frame during one time. The short-time average power level P of each time frame is computed, and the frequency-distribution function with which it is expressed as the number of the time frames corresponding to the numeric value of the average power level P is created.

[0047] Then, the increment (+1) of the counter a for the numbers of time frames is carried out. when the frequency distribution for the total frequency (the total number of time frames) defined beforehand can be created, it distinguishes whether it is way no and the frequency distribution for the total frequency are able to be created (it became $a>z$) Counter a is reset, and after detecting the 1st peak from the created frequency-distribution function and setting the average power level P_x of this 1st peak to Counter x , the created frequency-distribution function is cleared.

[0048] And each time frames DATA0-n are read one by one, and the average power level P of the read time frame distinguishes whether it is exceeding the average power level P_x of the noise section set to Counter x ($P>P_x$) from a frame memory.

[0049] In this case, when the average power level P of the read time frame exceeds the average power level P_x of the noise section (i.e., when the read time frame is not the noise section) ($P>P_x$), after

performing signal processing to read time frame DATA and considering as time frame DATA', D/A conversion of the data stream of that time frame is carried out, and it returns and outputs to an analog signal.

[0050] On the other hand, without performing signal processing to the read time frame, when the average power level P of the read time frame does not exceed the average power level P_x of the noise section (i.e., when the read time frame is the noise section) ($P \leq P_x$), D/A conversion of the data stream of the time frame is carried out as it is, and it returns and outputs to an analog signal.

[0051] That is, when the environmental sound signal S as shown, for example in drawing 7 is inputted, it outputs, after outputting as it is, carrying out analytic judgment to the voice section about the other time frame and performing signal processing to an environmental sound signal, without performing signal processing to an environmental sound signal about the time frame which attaches and shows O mark all over drawing by which analytic judgment was carried out to the noise section. It becomes possible to perform signal processing only to the signal of the time section of the voice contained in an environmental sound signal by this, and the output sound which is easy to catch can be obtained now.

[0052] Next, speech rate transform processing which applied this invention is explained with reference to drawing 8. In addition, since processing of the frequency-distribution clearance of the speech processing of drawing 6 is the same, this speech rate transform processing omits illustration and explanation. In this speech rate transform processing, each time frame is read one by one, and the average power level P of the read time frame distinguishes whether it is exceeding the average power level P_x of the noise section set to Counter x ($P > P_x$) from a frame memory.

[0053] Here, when the average power level P of the read time frame exceeds the average power level P_x of the noise section (i.e., when the read time frame is not the noise section) ($P > P_x$), wave expanding processing is performed to the data stream of the environmental sound signal of the time frame read (when it is the voice section). For example, as wave expanding processing is performed to the environmental sound signal which consists only of a sound signal as shown in drawing 9 (a) and it is shown in this drawing (b), on the whole, an environmental sound signal is elongated. And once storing the data stream after wave expanding in memory, it reads from memory one by one, D/A conversion is carried out, and it returns and outputs to an analog signal.

[0054] On the other hand, it is made not to output the time frame of the noise section to A/D-conversion processing as return and a result, without [when the average power level P of the time frame read from the frame memory does not exceed the average power level P_x of the noise section (i.e., when the read time frame is the noise section) ($P \leq P_x$), without it performs wave expanding processing to the read time frame, therefore] storing also in memory.

[0055] When the environmental sound signal which includes the noise as shown in drawing 10 (a) by this is inputted, the sound signal which this part was made to elongate can be inserted in by NEGUREKUTO [the noise section], and overflow of memory can be prevented, while being able to output now the sound signal which does not have an output sound lag so much to the sound signal inputted and becoming easy to catch an output sound.

[0056] In addition, although the frequency-distribution function which computes average power level (average strength) as description parameter value of an acoustical property, and is expressed with the number of time frames corresponding to average power level in the above-mentioned example was determined The description parameter value of the acoustical property of others, such as average power level, a peak factor, etc. after giving frequency weighting, is computed, and a frequency-distribution function can be determined based on the feature parameter of these acoustical properties so that clearly from the place already mentioned above.

[0057]

[Effect of the Invention] As explained above, according to the environmental sound analysis apparatus of claim 1, two or more time frames which divided the environmental sound signal into the signal for every predetermined time interval are generated. By computing the description parameter value of the generated acoustical property which was beforehand defined for every time frame, determining the frequency-distribution function showing the frequency distribution of the description parameter value of

an acoustical property, and detecting the peak part of this frequency-distribution function The voice section containing the time section which contains [from] a specific sound among environmental sounds, for example, the noise section which does not contain a conversation sound from the inside of an environmental sound, and a conversation sound can be extracted with high precision.

[0058] According to the environmental sound analysis apparatus of claim 2, it sets to the environmental sound analysis apparatus of above-mentioned claim 1. The A/D converter with which a frame division means carries out A/D conversion of the environmental sound signal, The frame memory which classified two or more time frames into two or more fields memorized, respectively, Since it has the division control means which makes each field of a frame memory memorize the conversion result of an A/D converter as a time frame for every part corresponding to the specific time die length which continued in time, division into the time frame of an environmental sound signal can be performed easily at high speed.

[0059] Since according to the environmental sound analysis apparatus of claim 3 it had a means to by_ which a parameter calculation means computed the average strength of a signal for every time frame, in above-mentioned claim 1 or the environmental sound analysis apparatus of 2 and has a means determine the frequency-distribution function with which a frequency-distribution function decision means is expressed with the number of the time frames corresponding to the numeric value of the average strength of a signal, the time section which contains [from] a specific sound among environmental sounds can extract with high precision.

[0060] Since it had the weighting means which carries out weighting based on a frequency to the signal of the time frame generated with the frame division means in the preceding paragraph of a parameter calculation means in above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 3 according to the environmental sound analysis apparatus of claim 4, the time section which contains [from] a specific sound among environmental sounds can be extracted much more with high precision.

[0061] Since according to the environmental sound analysis apparatus of claim 5 it had a means by which a parameter calculation means computed the peak factor of a signal for every time frame, in claim 1 or the environmental sound analysis apparatus of 2 and the frequency-distribution function decision means is equipped with a means to determine the frequency-distribution function expressed with the number of the time frames corresponding to the numeric value of the peak factor of a signal, the time section which contains [from] a specific sound among environmental sounds can be extracted with high precision.

[0062] Since it had a frame extract means to extract only the time frame corresponding to the peak part detected with the peak detection means in above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 5 according to the environmental sound analysis apparatus of claim 6, only the time section (signal part) which contains [from] a specific sound among environmental sound signals can be extracted.

[0063] According to the environmental sound analysis apparatus of claim 7, it sets to above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 6. Since it had a means to judge whether the location of the peak part of the determined frequency-distribution function is clear, and a means to change the total of the time frame used according to the judgment result of this means By the case where noise level is changed etc., when the location of the peak part of a frequency-distribution function is indefinite, by making [many] the total of the time frame to be used, a peak part can be detected more to high degree of accuracy, and analysis precision can be improved further.

[Translation done.]

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

TECHNICAL FIELD

[Industrial Application] This invention relates to the environmental sound analysis apparatus for extracting the time sections containing a specific sound, such as the section (silent section which does not contain a conversation sound) of the noise included in an environmental sound, and/or the section (voice section) of a conversation sound, concerning an environmental sound analysis apparatus.

[Translation done.]

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

PRIOR ART

[Description of the Prior Art] Since the environmental level or the environmental (operating environment) property of the noise which uses a mobile equipment is not fixed when using the device (a "mobile equipment" is called below.) which service spaces, such as a headphone stereo cassette tape recorder, hearing aid, a car stereo, a portable type telephone, and a walkie-talkie, do not specify generally, changing the acoustical property of a mobile equipment automatically according to an operating environment is performed.

[0003] If it does not change tone quality in coming to sense noisy and using it further in locations where noise level is still higher, such as inside of a subway, if it does not make sound volume low in using it to raising sound volume in the location where noise level, such as bustle, is high since the level of a conversation sound also becomes low in hearing aid in using it in quiet locations of Nighttime, such as domestic, it becomes impossible for example, to catch.

[0004] Then, in order to adjust the acoustical property of hearing aid automatically according to change of such an operating environment for example, the thing (the so-called ANS --) to which it was made to change automatically frequency characteristics (tone quality) and sound volume according to the sound pressure level of an input sound (environmental sound) or the sound pressure level, and its persistence time to ** hearing aid Refer to JP,52-50646,B and U.S. Pat. No. 4,025,721, ** The thing (refer to the so-called K amplifier and U.S. Pat. No. 5,131,046) to which it was made to change the gain of a high frequency band automatically according to the sound pressure level of the high frequency band of the input sound (environmental sound) to hearing aid like ANS is known.

[0005] Moreover, in hearing aid, in order that a user may make a conversation sound easy to catch, it is made to perform various signal processing to an input sound. For example, by performing analysis composition processing to an input sound, and changing a frequency to the lower one paying attention to being hard to catch a sound with a high frequency in many hearing impairment people A hearing impairment person tends to catch a conversation sound, and make him become, or What makes it easy to catch also to elderly people is performed by detecting, the break, i.e., the silent section, of a sentence of a conversation sound, securing time allowances by packing the time amount, and carrying out speech rate transform processing which makes late the conversation speed of sound outputted from hearing aid.

[0006] By the way, although it does not produce un-arranging so much even if it changes sound volume, tone quality, etc. of an output sound according to all environmental sound signals, if it is in a headphone stereo cassette tape recorder, a car stereo, etc. aiming at conversation in a mobile equipment which was mentioned above If it is in the hearing aid aiming at conversation, a portable type telephone, a walkie-talkie, etc., it is on the contrary hard coming to catch a conversation sound in having changed the acoustical property of an output sound according to all environmental sound signals. If a loudness level of sound is lowered according to all the inputted environmental sound signals, it becomes low to the level of not only the noise but a conversation sound, and it becomes impossible that is, to catch a conversation sound, since the conversation sound which contains its voice in an environmental sound is contained. [not only the noise but]

[0007] Moreover, it is on the contrary hard coming to catch a conversation sound in having performed

necessary signal processing to all environmental sound signals with the hearing aid which performs signal processing in order to make a conversation sound easy to catch. For example, if the frequency of the whole input sound is changed to the lower one, it changes also except a conversation sound, and a remarkable unnatural sound will hear or it will be hard coming to carry out identification of the sound source of an environmental sound. Moreover, although expanding processing of the environmental sound signal inputted must be carried out and it must once store in memory etc. in order to carry out speech rate transform processing, when it elongates to the signal of the noise, an output sound becomes redundancy, it becomes impossible to catch, and the memory space to need will increase.

[0008] Then, it discriminates from the noise and a conversation sound from among environmental sounds, and only the noise section or the voice section (section containing a conversation sound) is extracted from an environmental sound signal, and it is necessary to adjust an acoustical property or to perform necessary signal processing. Therefore, the section, the voice section, etc. of the noise which are conventionally included in an environmental sound, various things are proposed as an environmental sound analysis apparatus for extracting the time section containing a specific sound (JP,6-93199,B --) By comparing with the reference level as which, as for these, reference, such as JP,6-32001,B, determined the level of an environmental sound beforehand fundamentally, beyond over fixed time amount, for example, environmental sound level distinguishes the part at the time of below fixed level as it is the noise section.

[Translation done.]

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

EFFECT OF THE INVENTION

[Effect of the Invention] As explained above, according to the environmental sound analysis apparatus of claim 1, two or more time frames which divided the environmental sound signal into the signal for every predetermined time interval are generated. By computing the description parameter value of the generated acoustical property which was beforehand defined for every time frame, determining the frequency-distribution function showing the frequency distribution of the description parameter value of an acoustical property, and detecting the peak part of this frequency-distribution function The voice section containing the time section which contains [from] a specific sound among environmental sounds, for example, the noise section which does not contain a conversation sound from the inside of an environmental sound, and a conversation sound can be extracted with high precision.

[0058] According to the environmental sound analysis apparatus of claim 2, it sets to the environmental sound analysis apparatus of above-mentioned claim 1. The A/D converter with which a frame division means carries out A/D conversion of the environmental sound signal, The frame memory which classified two or more time frames into two or more fields memorized, respectively, Since it has the division control means which makes each field of a frame memory memorize the conversion result of an A/D converter as a time frame for every part corresponding to the specific time die length which continued in time, division into the time frame of an environmental sound signal can be performed easily at high speed.

[0059] Since according to the environmental sound analysis apparatus of claim 3 it had a means to by_ which a parameter calculation means computed the average strength of a signal for every time frame, in above-mentioned claim 1 or the environmental sound analysis apparatus of 2 and has a means determine the frequency-distribution function with which a frequency-distribution function decision means is expressed with the number of the time frames corresponding to the numeric value of the average strength of a signal, the time section which contains [from] a specific sound among environmental sounds can extract with high precision.

[0060] Since it had the weighting means which carries out weighting based on a frequency to the signal of the time frame generated with the frame division means in the preceding paragraph of a parameter calculation means in above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 3 according to the environmental sound analysis apparatus of claim 4, the time section which contains [from] a specific sound among environmental sounds can be extracted much more with high precision.

[0061] Since according to the environmental sound analysis apparatus of claim 5 it had a means by which a parameter calculation means computed the peak factor of a signal for every time frame, in claim 1 or the environmental sound analysis apparatus of 2 and the frequency-distribution function decision means is equipped with a means to determine the frequency-distribution function expressed with the number of the time frames corresponding to the numeric value of the peak factor of a signal, the time section which contains [from] a specific sound among environmental sounds can be extracted with high precision.

[0062] Since it had a frame extract means to extract only the time frame corresponding to the peak part detected with the peak detection means in above-mentioned claim 1 thru/or one environmental sound

analysis apparatus of 5 according to the environmental sound analysis apparatus of claim 6, only the time section (signal part) which contains [from] a specific sound among environmental sound signals can be extracted.

[0063] According to the environmental sound analysis apparatus of claim 7, it sets to above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 6. Since it had a means to judge whether the location of the peak part of the determined frequency-distribution function is clear, and a means to change the total of the time frame used according to the judgment result of this means By the case where noise level is changed etc., when the location of the peak part of a frequency-distribution function is indefinite, by making [many] the total of the time frame to be used, a peak part can be detected more to high degree of accuracy, and analysis precision can be improved further.

[Translation done.]

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

TECHNICAL PROBLEM

[Problem(s) to be Solved by the Invention] However, if it is in an environmental sound analysis apparatus [reference level / level / of an environmental sound] in order to extract the noise section and/or the voice section from an environmental sound signal, as mentioned above In noise level is high and being close to conversation sound level The section (noise section) when a conversation sound does not exist since an environmental sound does not become below fixed level even when a conversation sound does not exist is undetectable. Conversely, when conversation sound level is low the conversation sound exists -- environmental sound level -- below fixed -- becoming -- a conversation sound -- not existing (it being the noise section) -- the technical problem that the analysis precision of detect of an environmental sound is bad occurs.

[0010] This invention is made in view of the above-mentioned point, and it aims at offering the environmental sound analysis apparatus which can extract the time section when a specific sound is contained [from] among environmental sound signals with high precision.

[Translation done.]

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

MEANS

[Means for Solving the Problem] In order to solve the above-mentioned technical problem the environmental sound analysis apparatus of claim 1 the environmental sound analysis apparatus which extracts the time section when a specific sound is contained [from] among environmental sound signals -- it being and with a frame division means to generate two or more time frames which divided said environmental sound signal into the signal for every predetermined time interval A parameter calculation means to compute the description parameter value of the acoustical property which was generated with this frame division means and which was beforehand defined for said every time frame, It had a frequency-distribution function decision means to determine the frequency-distribution function showing the frequency distribution of said description parameter value computed with this parameter calculation means, and a peak detection means to detect the peak part of the frequency-distribution function determined with this frequency-distribution function decision means.

[0012] The environmental sound analysis apparatus of claim 2 is set to the environmental sound analysis apparatus of above-mentioned claim 1. The A/D converter with which said frame division means carries out A/D conversion of said environmental sound signal, The frame memory which classified said two or more time frames into two or more fields memorized, respectively, It considered as the configuration equipped with the division control means which makes each field of said frame memory memorize the conversion result of said A/D converter as said time frame for every part corresponding to the specific time die length which continued in time.

[0013] The environmental sound analysis apparatus of claim 3 was taken as the configuration which was equipped with a means by which said parameter calculation means computes the average strength of a signal for said every time frame, and is equipped with a means to determine the frequency-distribution function with which said frequency-distribution function decision means is expressed with the number of said time frames corresponding to the numeric value of the average strength of said signal in above-mentioned claim 1 or the environmental sound analysis apparatus of 2.

[0014] The environmental sound analysis apparatus of claim 4 was equipped with the weighting means which carries out weighting based on a frequency to the signal of said time frame generated with said frame division means in the preceding paragraph of said parameter calculation means in above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 3.

[0015] In claim 1 or the environmental sound analysis apparatus of 2, the environmental sound analysis apparatus of claim 5 was equipped with a means by which said parameter calculation means computes the peak factor of a signal for said every time frame, and was taken as the configuration which said frequency-distribution function decision means equips with a means to determine the frequency-distribution function expressed with the number of said time frames corresponding to the numeric value of the peak factor of said signal.

[0016] The environmental sound analysis apparatus of claim 6 was equipped with a frame extract means to extract only the time frame corresponding to the peak part detected with said peak detection means, in above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 5.

[0017] The environmental sound analysis apparatus of claim 7 was equipped with a means to judge

whether the location of the peak part of the frequency-distribution function which said frequency-distribution function decision means determined is clear, and a means to change the total of said time frame used according to the judgment result of this means, in above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 6.

[0018]

[Translation done.]

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

OPERATION

[Function] The environmental sound analysis apparatus of claim 1 generates two or more time frames which divided the environmental sound signal into the signal for every predetermined time interval with the frame division means. The description parameter value of the acoustical property which was generated with the parameter calculation means and which was beforehand defined for every time frame is computed. By determining the frequency-distribution function which expresses the frequency distribution of the description parameter value with a frequency-distribution function decision means, and detecting the peak part of a frequency-distribution function with a peak detection means The voice section containing the time section which contains [from] a specific sound among environmental sounds, for example, the noise section which does not contain a conversation sound from the inside of an environmental sound, and a conversation sound can be extracted with high precision.

[0019] The environmental sound analysis apparatus of claim 2 is set to the environmental sound analysis apparatus of above-mentioned claim 1. The A/D converter with which a frame division means carries out A/D conversion of the environmental sound signal, The frame memory which classified two or more time frames into two or more fields memorized, respectively, Since it has the division control means which makes each field of a frame memory memorize the conversion result of an A/D converter as a time frame for every part corresponding to the specific time die length which continued in time, division into the time frame of an environmental sound signal can be performed easily at high speed.

[0020] Since the environmental sound analysis apparatus of claim 3 was equipped with a means to by_ which a parameter calculation means computes the average strength of a signal for every time frame and is equipped with a means determine the frequency-distribution function with which a frequency-distribution function decision means is expressed with the number of the time frames corresponding to the numeric value of the average strength of a signal, it can extract the time section which contains [from] a specific sound among environmental sounds with high precision in above-mentioned claim 1 or the environmental sound analysis apparatus of 2.

[0021] In above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 3, since the environmental sound analysis apparatus of claim 4 was equipped with the weighting means which carries out weighting based on a frequency to the signal of the time frame generated with the frame division means in the preceding paragraph of a parameter calculation means, it can extract the time section which contains [from] a specific sound among environmental sounds much more with high precision.

[0022] In claim 1 or the environmental sound analysis apparatus of 2, since it had a means by which a parameter calculation means computed the peak factor of a signal for every time frame and the frequency-distribution function decision means is equipped with a means to determine the frequency-distribution function expressed with the number of the time frames corresponding to the numeric value of the peak factor of a signal, the environmental sound analysis apparatus of claim 5 can extract the time section which contains [from] a specific sound among environmental sounds with high precision.

[0023] In above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 5, since the environmental sound analysis apparatus of claim 6 was equipped with a frame extract means to extract

only the time frame corresponding to the peak part detected with the peak detection means, it can extract only the time section (signal part) which contains [from] a specific sound among environmental sound signals.

[0024] The environmental sound analysis apparatus of claim 7 is set to above-mentioned claim 1 thru/or one environmental sound analysis apparatus of 6. Since the frequency-distribution function decision means was equipped with a means to judge whether the location of the peak part of the determined frequency-distribution function is clear, and a means to change the total of the time frame used according to the judgment result of this means By the case where noise level is changed etc., when the location of the peak part of a frequency-distribution function is indefinite, a peak part can be detected more to high degree of accuracy by making [many] the total of the time frame to be used.

[Translation done.]

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

EXAMPLE

[Example] Hereafter, the example of this invention is explained based on an accompanying drawing. Drawing 1 is the block diagram showing one example of the environmental sound analysis apparatus concerning this invention.

[0026] This environmental sound analysis apparatus 1 inputs the signal which amplified the microphone signal from the microphone 2 which collects an environmental sound with amplifier 3 as an environmental sound signal S. A frame division means 4 to generate two or more time frames which divided the environmental sound signal S into the signal for every predetermined time interval, A parameter calculation means 5 to compute the description parameter value of the acoustical property which was generated with this frame division means 4 and which was beforehand defined for every time frame, A frequency-distribution function decision means 6 to determine the frequency-distribution function showing the frequency distribution of the description parameter value computed with this parameter calculation means 5, A peak detection means 7 to detect the peak part of the frequency-distribution function determined with this frequency-distribution function means 6, It has a frame extract means 8 to extract only the time frame corresponding to the peak part detected with this peak detection means 7, and an ambient noise analysis means 9 to analyze an ambient noise based on the time frame extracted with this frame extract means 8.

[0027] The frame division means 4 carries out A/D conversion of the environmental sound signal S, and consists of A/D converter 11 which carries out digital coding, a frame memory 12 which classified two or more time frames into two or more fields memorized, respectively, and a division control means 13 which makes each field of a frame memory 12 memorize the conversion result of A/D converter 11 as a time frame for every part corresponding to the specific time die length which continued in time.

[0028] The parameter calculation means 5 reads the time frame (environmental sound signal between specific time die length) memorized to each field of the frame memory 12 of the frame division means 4 one by one, and computes (average power level) as description parameter value in average strength for every time frame. The frequency-distribution function decision means 6 determines the frequency-distribution function showing the number of the time frames corresponding to the numeric value of the average power level of the time frame which the parameter calculation means 5 computed (frequency).

[0029] The peak detection means 7 detects the peak value which appears in the field where the number (frequency-distribution function) Kamihira ** power level of time frames corresponding to the numeric value of the average power level which the frequency-distribution function decision means 6 determined is low. The frame extract means 8 extracts from from only the time frame which has the average power level corresponding to the peak value detected with the peak detection means 7 among each time frame memorized by the frame memory 12, and is made to read it to the noise analysis means 9.

[0030] After the noise analysis means 9 carries out frequency analysis of the time frame (environmental sound signal of the specific time section) read from the frame memory 12 by the FFT algorithm, it computes each band power of a low frequency band, an inside frequency band, and a high-frequency band, and outputs the low frequency band power signal Pl, the inside frequency band power signal Pm, and the high-frequency band power signal Ph.

[0031] An operation of the example constituted as mentioned above is explained also with reference to drawing 2 thru/or drawing 5. If here explains the outline of the environmental sound analysis by this invention, it notes that the acoustical property of the sound which comes from the specific sound source in the scene of specification [this invention] is fixed in general. That is, since the description parameter value of acoustical properties, such as the level and frequency spectrum of the noise, and a peak factor, changes with the distance from the strength of a noise source, or a noise source, it cannot identify it with the absolute level. Moreover, the absolute level changes with an utterance sound, the reinforcement of utterance, and the distance from an utterance person also about a conversation sound.

[0032] However, in a specific scene, since the sound from a specific sound source is maintaining fixed level and a fixed acoustical property if a situation does not change, when it divides an environmental sound signal for every short time amount and asks for the frequency-distribution function of the description parameter value of the acoustical property in each time amount, in a frequency-distribution function, two or more peaks which embraced the class of sound will make it. For example, although a conversation sound is emitted on level somewhat higher than the noise when talking under the steady noise, as shown in a conversation sound at drawing 3, the silent section NB which is the section when a conversation sound is not surely contained exists, and although it is very a short time, since this silent section NB is the moment voice influence disappears, when an ambient noise exists in a perimeter, it turns into the section which the ambient noise itself appears.

[0033] At this time, what hits during conversation sound utterance among the environmental sound signals divided for every short time amount will have the average power corresponding to the level of a conversation sound, and what hits when it is not under utterance will have the average power corresponding to noise level. Therefore, two peaks will appear in a frequency-distribution function, the peak with larger average power will be equivalent to the signal at the time of being under utterance, and the peak of the smaller one will be equivalent to the signal at the time only of the noise which is not under utterance.

[0034] Then, what is necessary will be to perform signal-processing processing only to a signal with the average power of the peak of the larger one, performing the above-mentioned processing continuously that what is necessary is to collect only things with the average power equivalent to the peak of the smaller one of the divided environmental sound signals, and just to analyze if only the noise is analyzed, if signal-processing processing is performed only to a conversation sound.

[0035] Thus, even when the level of a conversation sound or the noise differs, by detecting the peak part of a frequency-distribution function, the time section containing a desired sound can be extracted with high precision, and necessary signal processing and adjustment of an acoustical property can be performed only in the required range.

[0036] When this example which analyzes an environmental sound below, detects the noise section and analyzes the acoustical property is explained concretely, the environmental sound analysis apparatus 1 the counter a for carrying out counting of the number of time frames which created frequency distribution as shown in drawing 2 -- resetting ($a=0$) -- the total number of time frames (this is called "the total frequency".) which is the need at creation of frequency distribution and which was defined beforehand It sets to Counter z (the z = total frequency). And A/D conversion of the environmental sound signal S which amplified the microphone signal from a microphone 1 with amplifier 2 is carried out with A/D converter 11 of the frame division means 4 for every fixed time amount, and it changes into the data stream of a digital sign.

[0037] Then, a break S, i.e., an environmental sound signal, is divided into the signal for every predetermined time interval with the number equivalent to the die length of the short time amount LD (sec) which defined the data stream of an environmental sound signal beforehand by the division control means 13 of the frame division means 4. The part (signal = data stream) corresponding to the die length of the time amount LD (sec) which continued in time is generated as one time frame, and the sequential storage of each generated time frame is carried out for every frame to each field which the frame memory 12 divided beforehand. namely, the signal part of the die length of the time amount LD (sec) short about the environmental sound signal S (for actually processing, it is a data stream after A/D

conversion) inputted as shown in drawing 4 -- as one time frame F -- time frames F1, F2, and F3 as -- it stores in each field of a frame memory 12. In addition, although the initiation timing of each time frame F is overlapped in the example of this drawing, this is for improving precision.

[0038] And the average power level P is computed for every time frame with the parameter calculation means 5 by reading the data stream of the environmental sound signal corresponding to the die length of the predetermined time amount LD memorized to each field of a frame memory 12 for every frame during one time. This average power level P can be computed by calculating $P = (x_0^2 + x_1^2 + x_2^2 + \dots + x_n^2) / n$, when the value acquired by f (Hz) and "LDxf" in a sampling frequency is set to n (measurement size).

[0039] Subsequently, based on the average power level P of the time frame computed with the parameter calculation means 5, a creation decision of the frequency-distribution function with which it is expressed as the number of the time frames corresponding to the numeric value of the average power level P is made with the frequency function decision means 6. That is, as shown in drawing 5, the frequency-distribution function which makes an axis of abscissa the average power level P, and makes an axis of ordinate the number of the time frames which have the average power level P concerned is created. This can form the counter corresponding to the average power level P of the time frame which the parameter calculation means 5 computed, and can obtain it by carrying out counting of the number of the time frames which have the average power level P concerned by carrying out the increment (+1) of the counter which corresponds whenever the power level P concerned is computed.

[0040] And by distinguishing whether after creating the frequency distribution about the time frame concerned read from the frame memory 12, the increment (+1) of the counter a was carried out, the value of Counters a and z was compared, and it became $a > z$, it judges whether the frequency distribution for the total frequency (the total number of time frames) defined beforehand have been created, and the above processings are repeated until it can create the frequency distribution for the total frequency.

[0041] Thus, a frequency-distribution function as shown in drawing 5 is obtained by creating the frequency distribution (function of the number of average power level pair time frames) of each average power level P about the time frame for the total frequency. namely, when the sound signal by the conversation sound mixes in an environmental sound signal Although the average power level P of the time frame which voice mixes becomes high, and it will become distribution as shown by the dotted line if crest type distribution as shown in the range of A in this drawing is shown and there is no ambient noise in the range of B of this drawing at this time When an ambient noise exists, as a continuous line shows, it is another crest (this is called "the 1st peak"). It appears.

[0042] Then, the peak detection means 7 detects the 1st peak (peak part of a frequency-distribution function) of the above from the frequency-distribution function created with the frequency-distribution function decision means 6, and it is presumed that the time frame which has the average power level P_x of this 1st peak is a time frame corresponding to an ambient noise. By it, the frame extract means 8 makes the time frame which has the average power level P_x , i.e., the time frame corresponding to an ambient noise, read from from among each time frame memorized by the frame memory 12, and is sent out for the noise analysis means 9. At this time, the frequency-distribution function decision means 6 clears the created frequency-distribution function.

[0043] In this case, when there are few time frames which have the average power level P_x and it is not suitable for analysis, it judges whether the location of the 1st peak is indefinite, and when the location of the 1st peak is indefinite, analysis precision can be secured by sending out the time frame which has the average power level before and behind the 1st peak for the noise analysis means 9. moreover, when such, it adjusts in the direction which increases the total frequency z (total of the time frame used for the decision of a frequency-distribution function) automatically -- if it is made like (z is changed), the precision of a frequency-distribution function can improve and analysis precision can also be raised further.

[0044] With the noise analysis means 9 which received the time frame corresponding to an ambient noise as mentioned above, after carrying out frequency analysis of the data stream of the time frame inputted by the FFT algorithm, each band power of a low frequency band, an inside frequency band, and

a high-frequency band is computed, and the low frequency band power signal P_l , the inside frequency band power signal P_m , and the high-frequency band power signal P_h are outputted.

[0045] Thus, two or more time frames which come to divide an environmental sound signal into the signal for every short time interval are generated. Compute the average power level (average strength) for every time frame as description parameter value, and by determining a frequency-distribution function based on the computed description parameter value, and detecting the peak part of this frequency-distribution function Only the signal of the time section (noise section) containing a necessary sound (the environmental sound which does not contain voice in this example = noise) can be extracted with high precision. The hearing aid sound of hearing aid can be adjusted now the optimal by adjusting acoustical properties, such as sound volume of hearing aid, and tone quality, using the low frequency band power signal P_l which followed, for example, was acquired as mentioned above, the inside frequency band power signal P_m , and the high-frequency band power signal P_h .

[0046] Next, the speech processing which applied this invention is explained with reference to drawing 6. This speech processing is made to perform signal processing only to parts other than the noise section of the environmental sound signals. First, while resetting the counter x for storing the average power level P_x of the time frame corresponding to the noise section ($x = 0$) After setting the total frequency which reset the counter a for the numbers of time frames like the above, and was beforehand set to the total counter z for frequencies, Sequential storing of the time frames DATA0-n which divided and generated similarly the environmental sound signal which amplified the microphone signal from a microphone with amplifier is carried out with the above-mentioned example having explained at a frame memory. The data stream of the time frames DATA0-n memorized to each field is read from a frame memory for every frame during one time. The short-time average power level P of each time frame is computed, and the frequency-distribution function with which it is expressed as the number of the time frames corresponding to the numeric value of the average power level P is created.

[0047] Then, the increment (+1) of the counter a for the numbers of time frames is carried out. when the frequency distribution for the total frequency (the total number of time frames) defined beforehand can be created, it distinguishes whether it is way no and the frequency distribution for the total frequency are able to be created (it became $a > z$) Counter a is reset, and after detecting the 1st peak from the created frequency-distribution function and setting the average power level P_x of this 1st peak to Counter x , the created frequency-distribution function is cleared.

[0048] And each time frames DATA0-n are read one by one, and the average power level P of the read time frame distinguishes whether it is exceeding the average power level P_x of the noise section set to Counter x ($P > P_x$) from a frame memory.

[0049] In this case, when the average power level P of the read time frame exceeds the average power level P_x of the noise section (i.e., when the read time frame is not the noise section) ($P > P_x$), after performing signal processing to read time frame DATA and considering as time frame DATA', D/A conversion of the data stream of that time frame is carried out, and it returns and outputs to an analog signal.

[0050] On the other hand, without performing signal processing to the read time frame, when the average power level P of the read time frame does not exceed the average power level P_x of the noise section (i.e., when the read time frame is the noise section) ($P \leq P_x$), D/A conversion of the data stream of the time frame is carried out as it is, and it returns and outputs to an analog signal.

[0051] That is, when the environmental sound signal S as shown, for example in drawing 7 is inputted, it outputs, after outputting as it is, carrying out analytic judgment to the voice section about the other time frame and performing signal processing to an environmental sound signal, without performing signal processing to an environmental sound signal about the time frame which attaches and shows O mark all over drawing by which analytic judgment was carried out to the noise section. It becomes possible to perform signal processing only to the signal of the time section of the voice contained in an environmental sound signal by this, and the output sound which is easy to catch can be obtained now.

[0052] Next, speech rate transform processing which applied this invention is explained with reference to drawing 8. In addition, since processing of the frequency-distribution clearance of the speech

processing of drawing 6 is the same, this speech rate transform processing omits illustration and explanation. In this speech rate transform processing, each time frame is read one by one, and the average power level P of the read time frame distinguishes whether it is exceeding the average power level P_x of the noise section set to Counter x ($P > P_x$) from a frame memory.

[0053] Here, when the average power level P of the read time frame exceeds the average power level P_x of the noise section (i.e., when the read time frame is not the noise section) ($P > P_x$), wave expanding processing is performed to the data stream of the environmental sound signal of the time frame read (when it is the voice section). For example, as wave expanding processing is performed to the environmental sound signal which consists only of a sound signal as shown in drawing 9 (a) and it is shown in this drawing (b), on the whole, an environmental sound signal is elongated. And once storing the data stream after wave expanding in memory, it reads from memory one by one, D/A conversion is carried out, and it returns and outputs to an analog signal.

[0054] On the other hand, it is made not to output the time frame of the noise section to A/D-conversion processing as return and a result, without [when the average power level P of the time frame read from the frame memory does not exceed the average power level P_x of the noise section (i.e., when the read time frame is the noise section) ($P \leq P_x$), without it performs wave expanding processing to the read time frame, therefore] storing also in memory.

[0055] When the environmental sound signal which includes the noise as shown in drawing 10 (a) by this is inputted, the sound signal which this part was made to elongate can be inserted in by NEGUREKUTO [the noise section], and overflow of memory can be prevented, while being able to output now the sound signal which does not have an output sound lag so much to the sound signal inputted and becoming easy to catch an output sound.

[0056] In addition, although the frequency-distribution function which computes average power level (average strength) as description parameter value of an acoustical property, and is expressed with the number of time frames corresponding to average power level in the above-mentioned example was determined The description parameter value of the acoustical property of others, such as average power level, a peak factor, etc. after giving frequency weighting, is computed, and a frequency-distribution function can be determined based on the feature parameter of these acoustical properties so that clearly from the place already mentioned above.

[Translation done.]

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] The functional block diagram showing one example of this invention

[Drawing 2] The outline flow Fig. of the environmental sound analysis processing with which operation explanation of drawing 1 is presented

[Drawing 3] The explanatory view of an input signal wave of the environmental sound with which explanation of environmental sound analysis is presented

[Drawing 4] The explanatory view with which explanation of frame division processing is presented

[Drawing 5] The explanatory view with which explanation of frequency-distribution function decision processing is presented

[Drawing 6] The flow Fig. of the example which applied this invention to speech processing

[Drawing 7] The explanatory view with which explanation of drawing 6 is presented

[Drawing 8] The flow Fig. of the example which applied this invention to speech rate transform processing

[Drawing 9] The explanatory view with which explanation of wave expanding processing of drawing 8 is presented

[Drawing 10] The explanatory view with which explanation of drawing 8 is presented

[Description of Notations]

1 [-- A frame division means, 5 / -- A parameter calculation means, 6 / -- A frequency-distribution function decision means, 7 / -- A frame extract means, 8 / -- A noise analysis means, 11 / -- An A/D converter, 12 / -- A frame memory, 13 / -- Division control means.] -- An environmental sound analysis apparatus, 2 -- A microphone, 3 -- Amplifier, 4

[Translation done.]

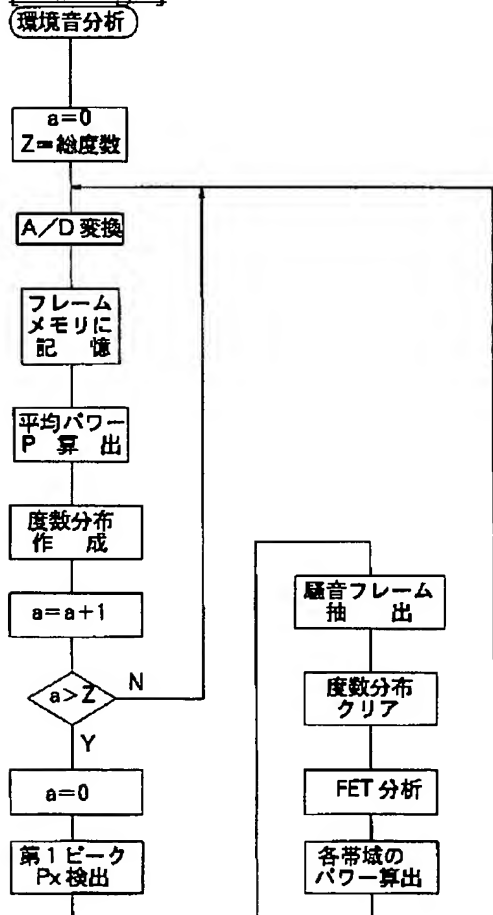
* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

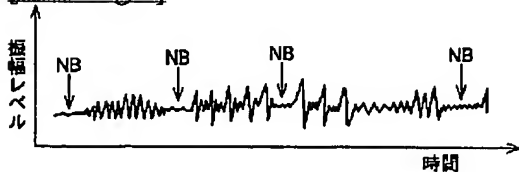
1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DRAWINGS

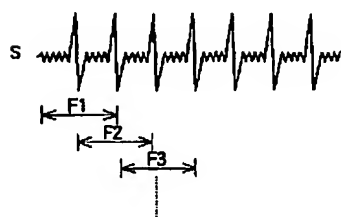
[Drawing 2]



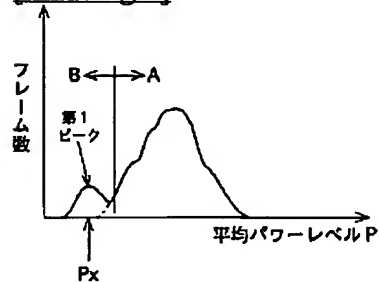
[Drawing 3]



[Drawing 4]



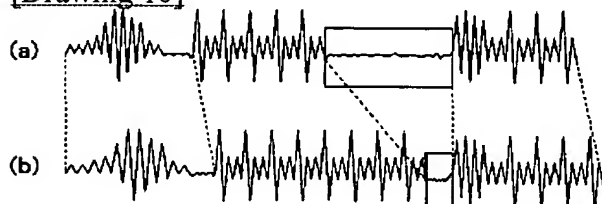
[Drawing 5]



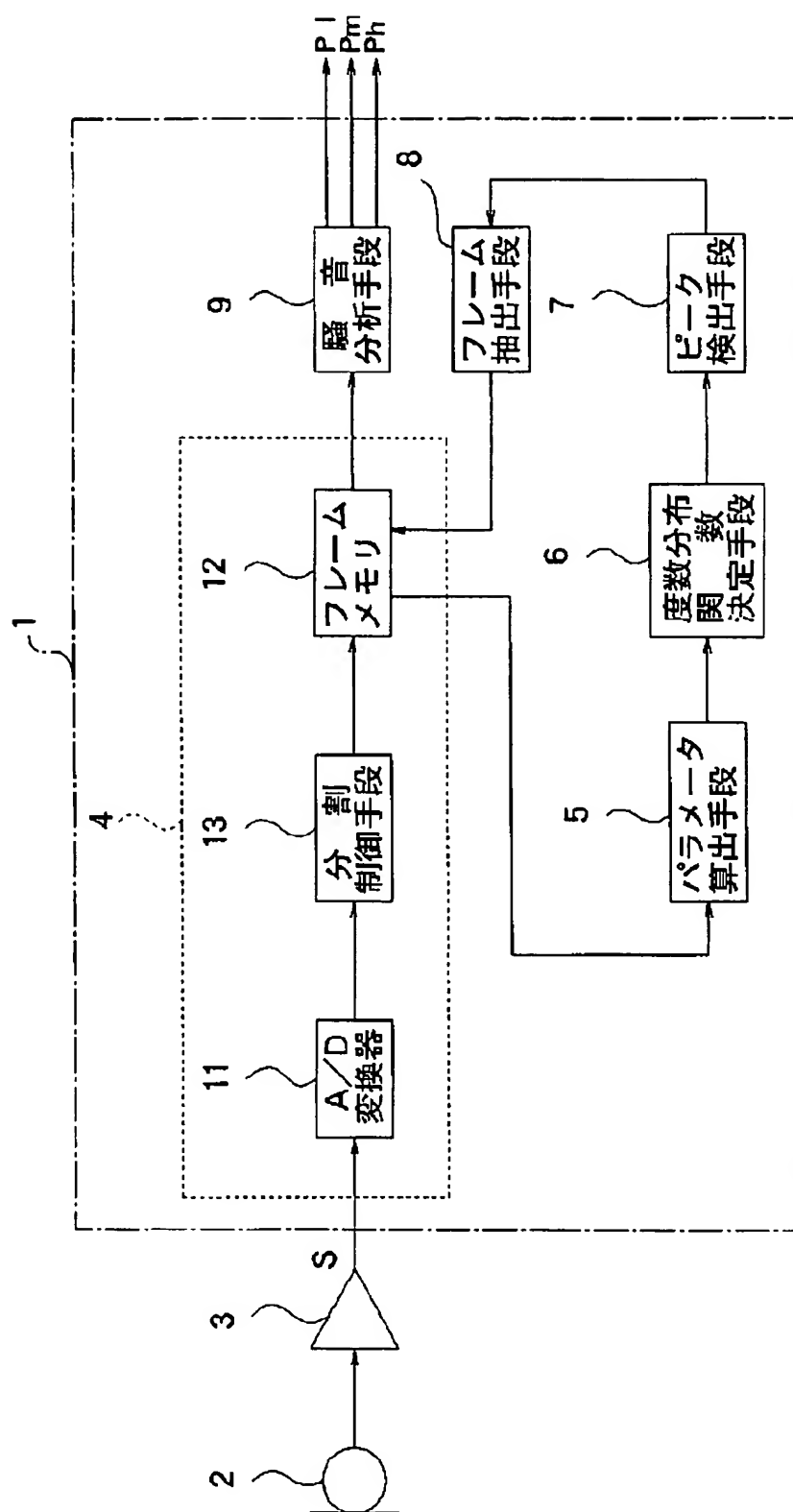
[Drawing 7]



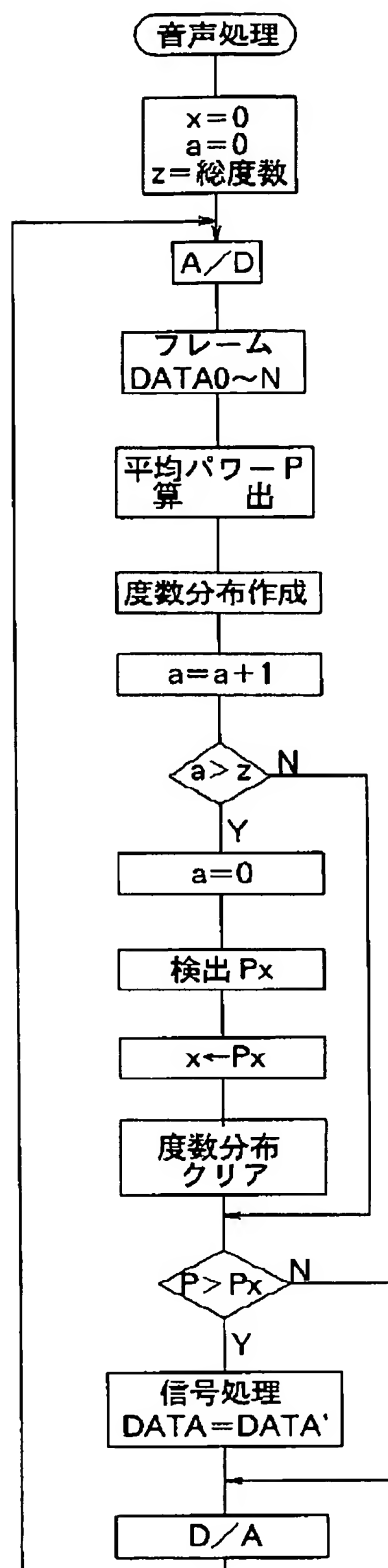
[Drawing 10]



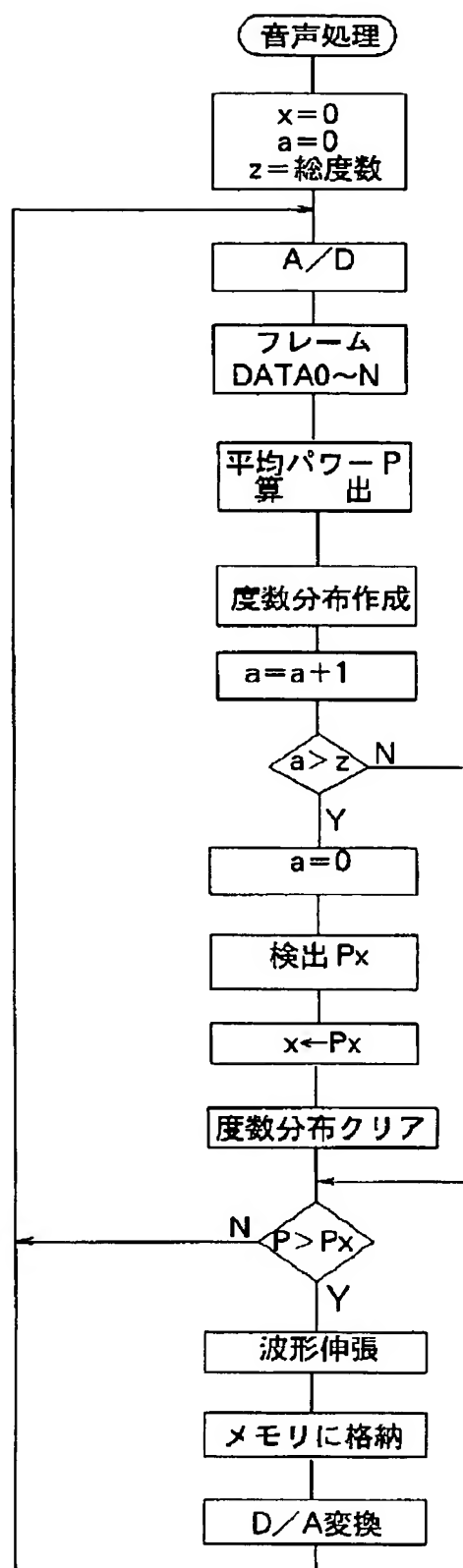
[Drawing 1]



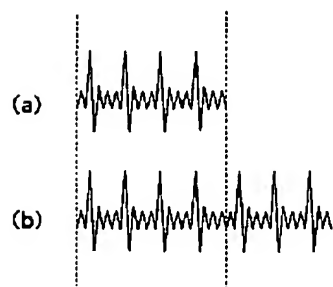
[Drawing 6]



[Drawing 8]



[Drawing 9]



[Translation done.]